VŠB - Technical University of Ostrava

Faculty of Economics

Department of Finance

# Predikce úpadku firem pomocí scoringového modelu

# Forecasting firm's bankruptcy by scoring model

Student: Bc. Yu Zhang

Supervisor of the diploma thesis: Ing. Jiří Valecký, Ph.D.

OSTRAVA 2016

VŠB - Technical University of Ostrava
Faculty of Economics
Department of Finance

# Diploma Thesis Assignment

| | |
|---|---|
| Student: | **Bc. Yu Zhang** |
| Study Programme: | N6202 Economic Policy and Administration |
| Study Branch: | 6202T010 Finance |
| Specialization: | 01 Finance |
| Title: | Predikce úpadku firem pomocí scoringového modelu<br>Forecasting Firm's Bankruptcy by Scoring Model |

Description:

1. Introduction
2. Modelling of Credit Risk
3. Description of Methodology Applied
4. Empirical Scoring Model
5. Conclusion

Bibliography
List of Abbreviations
Declaration of Utilization of Results from the Diploma Thesis
List of Annexes
Annexes

References:

CRAMER, Jan. *Logit models from economics and other fields.* Cambridge: Cambridge University Press, 2003. 173 p. ISBN 0-521-81588-6.
GOURIEROUX, Christian and Joann JASIAK. *Econometrics of individual risk: credit, insurance and marketing.* Princenton: Princeton University Press, 2007. 241 p. ISBN 978-0-691-12066-9.
HOSMER, David W. and Stanley LEMESHOW. *Applied logistic regression.* 2nd ed. New York: Wiley, 2000. 375 p. ISBN 0-471-35632-8.

Extent and terms of a thesis are specified in directions for its elaboration that are opened to the public on the web sites of the faculty.

Supervisor: **Ing. Jiří Valecký, Ph.D.**

Date of issue: 21.11.2014
Date of submission: 22.04.2016


Ing. Iveta Ratmanová, Ph.D.
*Head of Department*

prof. Dr. Ing. Dana Dluhošová
*Dean of Faculty*

The declaration

"Herewith I declare that I elaborated the entire thesis, including all annexes, independently. "

Ostrava dated 15.04.2016

Yu Zhang

Yu Zhang

# Contents

# 1. Introduction

Firms' bankruptcy is a proceeding of the firm which is unable to repay outstanding debts. At the same time, the poor clients, investors and creditors will not know the information as much as firms do. How to decrease the loss or how to avoid risk of invest on bad companies with a financial failure? Then, we need to study the reason of financial distress, find out the discipline and make prediction of firms.

Prediction model is almost popular in every field in daily life, especially in medicine, marketing research and economics, etc. It can be distinguished to rating and scoring. A rating is the evaluation or assessment of something, in terms of quality, quantity, or some combination of both. Scoring model is a different way of financial analysis based on quantitative of data submitted for the rated entity in previous economic period. Example is credit scoring – give the loan to those who with the higher scores.

The thesis is devoted to building a scoring model by using logit regression method to classify a company is a good performing or a bad one based on data of selected financial ratios of nearly 300 Czech firms.

There are five chapters in the thesis, list as follow: 1. Introduction; 2. modelling of Credit Risk; 3. Description of Methodology Applied; 4. Empirical Scoring Model; 5. Conclusion. At last, there are bibliography, list of Abbreviations and several annexes. And in details, chapter two is the general foundation description of credit risk modelling. And then more theory of method applied in the thesis will be discussed in chapter three. In chapter four, we will build the empirical scoring model with financial ratio data analysis, multiple regression analysis and show the model performance in ROC curves and this is also the most important part. At last, there will be the verification of model. Conclusion will be described in charter five.

# 2. Modelling of credit risk

Credit risk refers to the possibility that an unexpected change in counterparty's creditworthiness may generate a corresponding unexpected change in the market value of the associated credit exposure. There are three concepts: Default risk and migration risk; Risk as an unexpected event; Credit exposure (not just on-balance-sheet loans and securities).

One main challenge to credit risk management is the default event, which occurs if the debtor cannot meet its obligation decided by debt contract. This kind of default can be bonds default, corporate bankruptcy, mortgage foreclosure, the credit card charge-off, commodity trading default and so on. In this chapter we will present several approaches to measure credit risk.

## 2.1 The Regulatory Framework

In this section, we will introduce the standard approach of Basel framework for banking system. Risk in credit portfolios is the mix of specific risk in each individual security in the portfolio as risk factors. This method is easy to calculate but with few shortcomings like regardless of the functions of diversification effects.

The First Basel Accord of 1988, which is Basel I, makes the foundations for international minimum capital standard and banks became subject to regulatory capital requirements, coordinated by the Basel Committee on Banking Supervision. This Basel committee has been founded by the Central Bank Governors of the Group of Ten (G10) at the end of 1974. In the Central Bank Governors of the Group of Ten's view, the equity of the most important internationally active banks decreased to a dangerous level which leads to put forward of Basel I.

The downfall of Herstatt-Bank symbolized this concern which equity is used to absorb losses and to assure liquidity of a company or a bank. For the decrease insolvency risk of banks, losses and to minimize operating costs in the case of a

bankruptcy, Basel I targets to assure a suitable amount of equity and to create consistent international competitive conditions. The rules of the Basel Committee do not have legal force. The supervisory rules are rather intended to provide guidelines for the supervisory authorities of the individual nations can judge the rules and combine with individual situation or environment to applicate them.

Credit risk as the most important risk was the most important in the first Basel Accord. Within Basel I banks are supposed to keep at least 8% equity in relation to their assets. Now assets are weighted according to different degree of riskiness which are determined for four different borrower categories shown in Table 2.1.

Table 2.1 Risk weights for different borrower categories

| Risk Weight in % | 0 | 10 | 50 | 100 |
|---|---|---|---|---|
| Borrower Category | State | Bank | Mortgages | Companies and Retail Customers |

The required equity can then be computed as

$$\text{Minimal Capital} = \text{Risk Weighted Assets} \times 8\%. \tag{2.1}$$

Hence the portfolio credit risk is measured as the sum of risk weighted assets that are the sum of different risk weights assets distinguish by borrower categories and goodwill. Since this approach did not take into account of market risk, in 1996 an modification to Basel I has been published which allows for both a standardized approach and a method based on internal Value-at-Risk (VaR) models for market risk in larger banks or corporates. But the main criticism of Basel I is still exists. Which is we have mention before Basel I does not account for methods to decrease risk like through portfolio diversification. What's more, the approach measures risk in an insufficiently differentiated way since minimal capital requirements are computed independent of the borrower's goodwill. These drawbacks lead to the development of the Second Basel Accord from 2001 onwards. In June 2004 the Basel Committee on Banking Supervision released a Revised Framework or Basel II. The main targets of

Basel II are the same as in Basel I as well. And Basel II focuses not only on market and credit risk but also puts operational risk on the agenda.

Basel II is structured in a three-pillar framework. Pillar 1 builds up the details for adopting more risk sensitive minimal capital requirements, so-called regulatory capital for banking organizations, Pillar 2 published principles for the supervisory review process of capital adequacy and Pillar 3 contributes to establish market discipline by enhancing transparency in banks' financial reporting.

A new risk category or a basic innovation of Basel II was the creation of a new risk category-consideration of operational risk, and taken it into account in the new accord.

Basel is structured in a three-pillar framework. Pillar 1 builds up the details for adopting more risk sensitive minimal capital requirements, so-called regulatory capital for banking organizations, Pillar 2 published principles for the supervisory review process of capital adequacy and Pillar 3 contributes to establish market discipline by enhancing transparency in banks' financial reporting.[1]

## 2.2 Credit ratings and credit scoring

Here we study the probability of default from the actual credit performance which include two aspects:

- Credit rating: An assessment of the credit worthiness of a borrower in general terms or with respect to a particular debt or financial obligation. Credit assessment and evaluation for companies and governments is generally done by a credit rating agency such as Standard & Poor's or Moody's. These rating agencies are paid by the entity that is seeking a credit rating for itself or for one of its debt issues.
- Credit scoring: A statistically derived numeric expression of creditworthiness of

[1] See in Bibliography [1]

study object that is used by lenders to access the likelihood that a person will repay his or her debts.

"The big three" in credit rating agency are Standard & Poor's, Moody's and Fitch Ratings which controls approximately 95% of credit rating business. As shown in Table 2.2, Aaa, Aa, A, Baa, Ba, B are used to describe the likelihood of default, the higher the rating, the better investment grade or the bond is. Low rating means higher probability of default. A high credit score indicates a stronger credit profile and will generally result in lower interest rates charged by lenders.

Table 2.2: Bond credit rating table

| | Moody's | S&P | Fitch | Meaning |
|---|---|---|---|---|
| Investment Grade | Aaa | AAA | AAA | Prime |
| | Aa1 | AA+ | AA+ | High Grade |
| | Aa2 | AA | AA | |
| | Aa3 | AA- | AA- | |
| | A1 | A+ | A+ | Upper Medium Grade |
| | A2 | A | A | |
| | A3 | A- | A- | |
| | Baa1 | BBB+ | BBB+ | Lower Medium Grade |
| | Baa2 | BBB | BBB | |
| | Baa3 | BBB- | BBB- | |
| Junk | Ba1 | BB+ | BB+ | Non-Investment Grade Speculative |
| | Ba2 | BB | BB | |
| | Ba3 | BB- | BB- | |
| | B1 | B+ | B+ | Highly Speculative |
| | B2 | B | B | |
| | B3 | B- | B- | |
| | Caa1 | CCC+ | CCC+ | Substantial Risk |
| | …D | …D | …DDD | In Default |

Here is a credit ratings table which shows each rating level from the three major credit rating agencies, and a brief explanation of what each level means.

For example, U.S. treasury bonds have the highest credit rating, because of its never default and stable interest. Vice versa, high-yield (junk) bonds are rated much lower （receive higher interest）

In credit scoring, the most familiar to consumers will be FICO, developed by Fair Isaac Corporation. It ranges from 300 (very poor) to 850 (best), and intends to represent the creditworthiness of a borrower such that he or she will repay the debt. For the same borrower, the three major U.S. credit bureaus often report inconsistent FICO scores based on their own proprietary models.

Table 2.3：Credit score table

| | |
|---|---|
| 760-850 | EXCELLENT |
| 700-759 | VERY GOOD |
| 723 | MEDIAN FICO SCORE |
| 660-699 | GOOD |
| 687 | AVERAGE FICO SCORE |
| 620-659 | NOT GOOD |
| 580-619 | POOR |
| 500-579 | VERY POOR |

Here is a credit score table which shows each score level from FICO, and a brief explanation of what each level means.

## 2.3 Expected and Unexpected Loss

It is hard to forecast the losses of for example a bank will suffer in a certain period of time, but we can still predict the average level of credit loss. These are expected loss (EL). Suppose we have N number of obligors and use n as the quantities specific to obligor n. The expected loss function is as follow，

$$EL_n = EAD_n * LGD_n * PD_n \quad (2.2)$$

Where EAD is exposure at default, amount to which the bank was exposed to borrower at the time of default, measured in currency; LGD is loss given default, magnitude of likely loss on the exposure, expressed as a percentage of the exposure; PD is probability of default of borrower

When peak losses exceed our expected level, these large losses are so-called unexpected losses (UL). The unexpected loss can be defined as the variability of the loss around its mean value, i.e. around the EL. It is calculated as a standard deviation from the mean of the loss variable at a certain confidence level. Unlike expected loss, the unexpected loss of a portfolio is not calculated by adding the unexpected loss of individual assets. This is because unless there is perfect correlation, the standard deviation of sum will not be the same as the sum of standard deviation. The formula is:

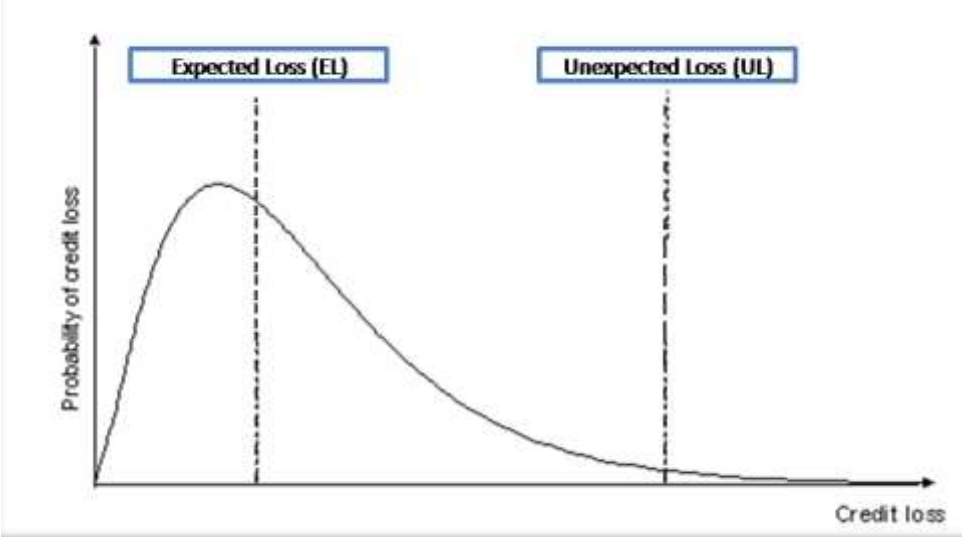
Peak losses exceed our expected level, these large losses are so-called unexpected losses (UL). The unexpected loss can be defined as the variability of the loss around its mean value, i.e. around the EL. It is calculated as a standard deviation from the mean of the loss variable at a certain confidence level. Unlike expected loss, the unexpected loss of a portfolio is not calculated by adding the unexpected loss of individual assets. This is because unless there is perfect correlation, the standard deviation of sum will not be the same as the sum of standard deviation. The formula is:

$$UL_n = \sqrt{V[L_n]} = \sqrt{V[EAD_n * LGD_n * D_n]} \quad (2.3)$$

In case the default indicator $D_n$ and the LGD variable are uncorrelated (and the EAD is constant), the UL on borrower n is given by:

$$UL_n = EAD_n \sqrt{VLGD_n^2 * PD_n + ELGD_n^2 * PD_n(1 - PD_n)} \quad (2.4)$$

Figure 2.1: Expected loss and unexpected loss[2]




As shown in the Figure 2.1, we can forecast the expected loss for example we know some business cost or liquidation cost (risk of default) which can be calculated. Usually unexpected loss usually got higher potential credit losses.

## 2.4 Merton model

Credit risk models can be divided into two fundamental classes of models, which are structure or asset-value models, and reduced-form or default rate models.

Merton model, also referred to as "Asset Value Model" is a model named after Robert C. Merton and is used to evaluate the credit risk of a corporation's debt. More specifically, we employ the model to determine a company's ability to service its debt, meet its financial obligations and to gauge the overall possibility of credit default.

Generally, the value of the firm's assets is assumed to obey a lognormal spread process with a constant volatility. There are only two classes of securities: equity and debt. The Merton model assumes the asset value of the company follows some random steps $(V_t)_{t\geq 0}$. The equity receives no dividends and the company's debt is given with face value $B$ that will become due at time $T$ in the future. In the Merton model default can occur only at the maturity $T$ of the bond. Denote the value at time $t$

[2] See in Electronic reference [12]

of equity and debt by $S_t$ and $B_t$. In a frictionless market with no taxes or transaction costs, the value of the firm's assets is given by the sum of debt and equity, for example，

When peak losses exceed our expected level, these large losses are so-called unexpected losses (UL). The unexpected loss can be defined as the variability of the loss around its mean value, i.e. around the EL. It is calculated as a standard deviation from the mean of the loss variable at a certain confidence level. Unlike expected loss, the unexpected loss of a portfolio is not calculated by adding the unexpected loss of individual assets. This is because unless there is perfect correlation, the standard deviation of sum will not be the same as the sum of standard deviation. The formula is:

$$V_t = S_t + B_t, 0 \leq t \leq T \tag{2.5}$$

There are two possibilities at maturity:

- $V_T > B$: the value of the company's assets is bigger than the debt. In this situation, debtholders receive $B_T = B$, shareholders receive residual value which is $S_T = V_T - B$ with no default.
- $V_T \leq B$: The value of the company's assets is less than the debt which means the company is unable to meet its financial obligations and defaults. Debtholders take ownership of the firm and shareholders have nothing, which is $B_T = V_T$, $S_T = 0$.

To sum up, the payment to the shareholders at time T is calculated as follows:

$$S_t = \max(V_T - B, 0) = (V_T - B)^+ \tag{2.6}$$

And debtholders receive:

$$B_T = \min(V_T, B) = B - (B - V_T)^T \tag{2.7}$$

According to put-call parity, the firm's debt consist of a risk-free bond that guarantees payment of $B$ plus a short European put option on the firm's assets with exercise price equal to the promised debt payment $B$. we assume that the security

price follows a geometric Brownian motion,

$$dV_T = \mu v\, V_t dt + \sigma v\, V_t dWt,\ 0 \le t \le T \tag{2.8}$$

Where constant $\mu v \in R$, $\sigma v$ is above zero, and the standard Brownian motion $(Wt)_t \geq 0$. Then initial value $V_0$ is calculated as follows，

$$V_T = V_0 \exp\left(\left(\mu v - \frac{1}{2}\sigma_v^2\right)T + \sigma v W_T\right. \tag{2.9}$$

This can be transformed to:

$$\ln V_T \sim N\left(lnV_0 + \left(\mu v - \frac{1}{2}\sigma_v^2\right)T, \sigma_v^2 T\right. \tag{2.10}$$

Based on European call option principle and risk neutral pricing theory, we can get:

$$S_t = V_t * \Phi\left(d_{t,1}\right) - B * e^{-r(T-t)} * \Phi(d_{t,2}) \tag{2.11}$$

Here r refers to constant risk-fee interest rate. $d_{t,1} = \frac{\ln\left(\frac{V_t}{B}\right)+(r+\frac{\sigma_V^2}{2})(T-t)}{\sigma_V\sqrt{T-t}}$ and $d_{t,2} = d_{t,1} - \sigma v\sqrt{T-t}$.

The default probability of the firm by time $T$ is the probability that the shareholders of the firm will not exercise their call option to buy the assets of the company for $B$ at time $T$. the profitability can be calculated,

$$P(V_T \le B) = P(\ln V_T \le \ln B) = \Phi\left(\frac{\ln(B/V_0) - \left(\mu V - \frac{\sigma_V^2}{2}\right)T}{\sigma v\sqrt{T}}\right. \tag{2.12}$$

As shown in formula (2.11), we can see that the profitability of default is increasing as B increasing and decreasing in $V_0$ and $\sigma v$.

## 2.5 Credit-Scoring Models

Multivariate models which use the main economic and financial indicators of a company as input, attributing a weight to each of them, that reflects its relative importance in forecasting default. The result is an index of creditworthiness expressed as a numerical score, which indirectly measures the borrower's probability of default.

- Three type of models will be presented:

- Linear discriminant analysis;
- Regression models (linear, logit and probit regression);
- Some recent heuristic inductive models such as neural networks and genetic algorithms.

### 2.5.1 Linear discriminant analysis

Linear discriminant analysis (LDA) is a generalization of Fisher's linear discriminant, a method used in statistics and machine learning to find a linear combination of features that characterizes or separates two or more classes of objects or events.

LDA is also closely related to principal component analysis and factor analysis in that they both look for linear combinations of variables which best explain the data. LDA explicitly attempts to model the difference between the classes of data. PCA on the other hand does not take into account any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities. Discriminant analysis is also different from factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables (also called criterion variables) must be made.

LDA works when the measurements made on independent variables for each observation are continuous quantities. When dealing with categorical independent variables, the equivalent technique is discriminant correspondence analysis.

For a given borrower i, we calculate the score $z$ as follows:

$$z_i = \sum_{j=1}^{n} \gamma_j x_{i,j} \tag{2.13}$$

Where $x$ represents variables usually financial indicators, $\gamma$ is its coefficient within the estimated model and n is the number of indicators.

The vector of gamma coefficients in formula (3.4) is calculated as follows, for more detail see Resti and Sironi (2007):

$$\alpha = \frac{1}{2}\gamma'(x_A + x_B) \tag{2.14}$$

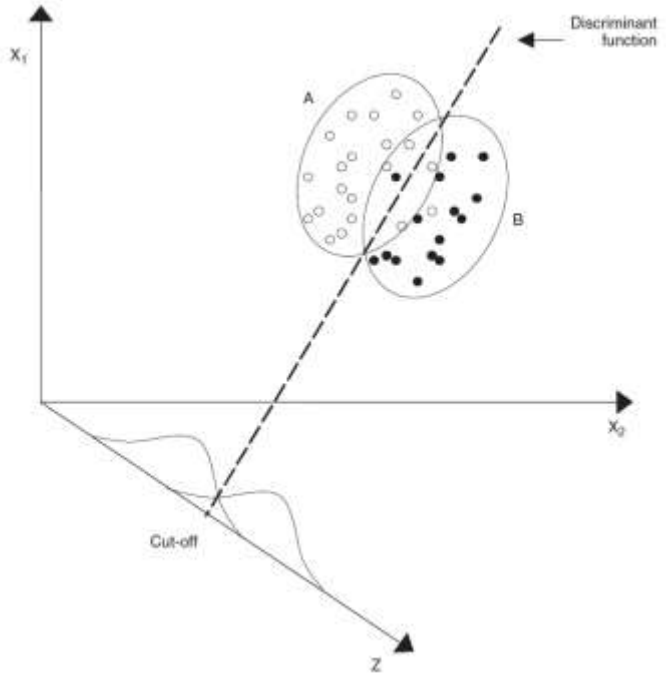
$$\gamma = \sum^{-1}(x_A - x_B) \tag{2.15}$$

Where $x_A$ and $x_B$ are vectors include mean value of independent variables.

Linear discriminant analysis can be used to produce a direct estimate of the probability of default. As shown in Altman (1968) or Resti and Sironi (2007) for more detail, that the company's probability of default can be calculated:

$$\text{PD} = \text{p}(B|x_i) = \frac{1}{1+\frac{1-\pi_B}{\pi_B} * e^{z_i - \alpha}} \tag{2.16}$$

Figure 2.2: Linear discriminant function[3]




In bankruptcy prediction based on accounting ratios and other financial variables, linear discriminant analysis was the first statistical method applied to systematically explain which firms entered bankruptcy vs. survived.

The best-known discriminant score applied to credit risk is probably Z-score model the one developed by Edward Altman in 1968. It is a function of five independent variables, which is still a leading model in practical applications and the formula is as follows:

[3] See in Bibliography [8]

$$z_i = 1.2 * x_{i,1} + 1.4 * x_{i,2} + 3.3 * x_{i,3} + 0.6 * x_{i,4} + 1.0 * x_{i,5} \quad (2.17)$$

Where: $x_1$ refers to working capital divided by total assets, a measure of the net liquid assets of the firm relative to the total capitalization;

$x_2$ means retained profits divided by total assets, the account which reports the total amount of reinvested earnings and/or losses of a firm over its entire life and also measures the leverage of a firm;

$x_3$ is earnings before interest and tax divided by total assets, which measures the true productivity of the firm's assets, independent of any tax or leverage factors;

$x_4$ denotes market value of equity divided by book value of total liabilities, which shows how much the firm's assets can decline in value before the liabilities exceed the assets and the firm becomes insolvent;

$x_5$ is turnover divided by total assets and measures of the management's capacity in dealing with competitive conditions.

If Z＜1.8, area of perdition of the firm's bankruptcy

If 1.8＜Z＜2.99, the grew zone

If Z＞2.99, prediction of non-bankruptcy area (safe area)

This model is to estimate the operation situation of companies by function constructed of selected five financial ratios.

We have three questions to build up this model, 1) which ratio is the most important in detecting bankruptcy potential. 2) What weights should be attached to those selected ratios, and 3) how to objectively established the weights.

When utilizing a comprehensive list of financial ratios in assessing a firm's bankruptcy, it's easy to find out that some of the measurements will have a high degree of correlation or collinearly of each other.

The MDA technique has the advantage of considering an entire profile of characteristics. To help us analyzing the entire variable profile of the object simultaneously rather than just examining its individual characteristics

### 2.5.2 Regression models – Probit model

Logistic regression model which we applied in this these will be introduced in next chapter. Probit regression analysis is estimate the probability that an event occurs or not like married or not married. The purpose of the model is to estimate the probability that an observation with particular characteristics will fall into a specific one of the categories. The response $y_i$ is equals to zero or one (occurs or not occurs)

$$P_i = f(\alpha + \beta' * x_i) \tag{2.18}$$

Where $x_i$ are financial indicators or ratios, $\alpha$, $\beta$ are estimate parameters.

Calculated by using cumulative distribution function of normal distribution:

$$P_i = \int_{-\infty}^{\alpha+\beta'*x_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt \tag{2.19}$$

According to the features of nonlinear it is necessary to use maximum likelihood method for parameters estimation. Given $P_i$ and assuming that defaults are independent, we can form the logarithm of likelihood function as follows:

$$\ln L = \sum_{i=1}^{n} y_i ln Pi + \sum_{i=1}^{n} (1 - y_i) \ln(1 - P_i) \tag{2.20}$$

# 3 Description of Methodology Applied

The discipline of prediction of firms' bankruptcy is to pick up appropriate financial ratios as independent variables to estimate the probability of bankruptcy. In this chapter, we'll introduce financial ratios at first and then talk about the building logistic regression model and ROC curve. At last there will be the verification model using residual analysis and different data.

## 3.1 Description of data—financial ratios

**Data:** We will use sample of 372 Czech companies for estimation--172 bankruptcy firms and 200 non-bankruptcy firms. Then we have 20 selected variables which is financial ratios (profitability ratios, solvency ratios…) at the beginning from the financial statements.

**Financial ratios:**

**F**inancial ratios are used to try to evaluate the overall financial condition of a corporation or other organization. Values used in calculating financial ratios are taken from the balance sheet, income statement and cash flow statement or (sometimes) the statement of retained earnings. All these comprise the firm's "accounting statements" or financial statements. The statements' data is based on the accounting method and accounting standards used by the organization.

Generally, each kind of financial ratios reflect to the financial situation of company more or less. But the information these ratios give us are overlapping to some extent, which means these ratio have a high relevance or substitutability, and influence the accuracy and efficiency of assessment of our model.

Financial crisis usually causes credit crisis, which leads companies or other investors make different choices. We can find out some special financial signs which can show borrower's financial position, it can helps firms determine their credit rating,

provide the evidence for credit and investment. Due to this motivation, financial analyses of corporate become important an important measurement for judge the credit risk of firms. Except studying company's financial ratios to measure the probability of default from the borrowers, we also need to take basic financial ratios analysis into account for better credit risk control.

Here, let's just focus on financial ratios and to introduce method of choosing independent variables later in Chapter 3.2.

***ROE***

Return on equity is the rate of return on the shareholders' equity of the common stock owners. So here we use the net income to measure a company's efficiency at generating profits from every unit of shareholders' equity. We can calculate it like:

$$ROE = \frac{net\ profit}{equity} \tag{3.1}$$

***ROA (EBITDA/A)***

Return on assets is calculated by dividing a company's annual earnings by its total assets, shows how efficient management is at using its assets to generate earnings. We can calculate it like:

$$ROA = \frac{operating\ profit + depreciation\ and\ ammortisation}{assets} \tag{3.2}$$

***ROA (EBIT/A)***

Simply calculation:

$$ROA = \frac{operating\ profit}{assets} \tag{3.3}$$

***ROS (EBIT/S)***

A ratio widely used to evaluate a company's operational efficiency. It is calculated using this formula:

$$ROS = \frac{EAT}{sales} \quad (3.4)$$

***D/A***

Total debt to assets ratio is a leverage ratio that defines the total amount of debt relative to assets. The higher the ratio is, the higher the degree of leverage and also financial risk is. The formula is:

$$ROS = \frac{total\ debt}{total\ assets} \quad (3.5)$$

***BANK LOAN/ASSETS***

Bank loan to assets ratio is calculating the percentage of a company's assets that are provided via bank loan. Companies with high bank loan/asset ratios are said to be "highly leveraged," not highly liquid as stated above. The formula is as follows:

$$Bank\ loan\ to\ assets = \frac{bank\ loan}{assets} \quad (3.6)$$

***DSI***

Day's sale of inventory is financial measure of a company's performance that shows how long it takes a company to turn its inventory into sales. Generally, the lower the DSI is the better. Here is how the DSI is calculated:

$$DSI = \frac{inventory*360}{sales} \quad (3.7)$$

***DSO***

Days sales outstanding in accounts receivables is a measure of the average number of days that a company takes to collect revenue after a sale has been made. The lower the ratio is, the less the days take to collect its accounts receivable. The formula is as follows:

$$DSO = \frac{accounts\ receivable*360}{sales} \quad (3.8)$$

***A*360/Sales***

Day's sales in assets ratio is similar to asset turnover ratio indicates that how quickly a company can turn an asset into cash on average. The formula is as follows:

$$Day's\ sales\ in\ assets = \frac{assets*360}{sales} \tag{3.9}$$

***CURRENT RATIO***

Current ratio is a ratio of current assets divided by current liabilities. It is a financial ratio that measures whether or not a firm has enough resources to pay its debts over the next 12 months. The current ratio indicates a firm's market liquidity and ability to meet creditor's demands. In general it's between 1.5 and 3 for healthy businesses. If it is below 1, the company's current liabilities are more than current assets, which means the company may have problems facing its short-term obligations. If the current ratio is higher than 3, which means current assets are much more than current liabilities, the company didn't use its money well or efficient. We can calculate it like:

$$current\ ratio = \frac{current\ assets}{current\ liability} \tag{3.10}$$

***QUICK RATIO***

Quick ratio is a ratio of quick assets divided by current liabilities. Quick assets include those current assets that can be quickly turned to cash at close to their face values, which equals current assets minus inventory. Generally, the ratio should be 1:1 or higher and the higher the ratio, the greater the company's liquidity (i.e., the better able to meet current obligations using liquid assets). We can calculate it like:

$$quick\ ratio = \frac{current\ assets - inventory}{current\ liability} \tag{3.11}$$

***CASH RATIO***

Cash ratio is to measure the ability of a company to facing current obligations by just using cash and cash equivalents and short-term investment. The formula is as follows: capital employed/Fixed assets

$$cash\ ratio = \frac{cash\ and\ cash\ equivalent + short-term\ investment}{cash\ liability} \quad (3.12)$$

***Working capital to current assets***

The Working Capital to current assets ratio measures a company's ability to cover its short term financial obligations which is current liability by comparing its remaining liquid assets to the total current assets. The formula is:

$$working\ capital\ to\ assets = \frac{current\ assets - current\ liability}{current\ assets} \quad (3.13)$$

***Gross profit margin***

Gross profit margin is one of the profitability ratios. The higher the ratios is, the better the competition position the company is .The formula is:

$$gross\ profit\ margin = \frac{gross\ margin}{sales} \quad (3.14)$$

***Financial leverage***

Financial leverage is ratio of total assets divided by total shareholder's equity. It can be described as the extent to which a business or investor is using the borrowed money. The formula is as follows:

$$financial\ leverage = \frac{total\ assets}{equity} \quad (3.15)$$

***Interests/ EBIT***

Interests/EBIT is opposite of interest coverage ratio is one of the key financial ratios used in assessing the credit of a corporation which shows the pre-interest earnings to the charges required on debt. The calculation is:

$$interests/EBIT\ ratio = \frac{intereset}{EBIT} \quad (3.16)$$

***Capital employed/Fixed assets***

Capital employed/fixed assets ratio indicate the extent to which the long term funds are sunk in fixed assets which are supplied by creditors and owners of the firm.

The ratio is calculated by:

$$capital\ employed\ ratio = \frac{capital\ employed\ ratio}{fixed\ assets} \quad (3.17)$$

***Current liabilities/Asset***

Current liabilities to assets ratio is a solvency ratio to examine how much of a company's assets is made of liabilities. The calculation is:

$$CL\ to\ assets = \frac{current\ liabilities}{assets} \quad (3.18)$$

***D/E***

Debt to equity ratio indicates the proportion of equity and debt the company is using to finance its assets.

$$Debt\ to\ equity = \frac{total\ liabilities}{equity} \quad (3.19)$$

***Bank loans/Equity***

Bank loan to equity ratio is calculating the percentage of a company's equity that is provided via bank loan. The calculation is:

$$bank\ loans\ to\ equiy = \frac{bank\ loans}{equity} \quad (3.20)$$

## 3.2 Univariate analysis of variables

Univariate analysis is the simplest form of quantitative (statistical) analysis. The analysis is carried out with the description of a single variable in terms of the applicable unit of analysis. For example, if the variable "age" was the subject of the analysis, the researcher would look at how many subjects fall into given age attribute categories.

A basic way of presenting univariate data is to create a frequency distribution of

the individual cases, which is to present the number of cases in the sample that belong to each category of values of the variables. This can be done in a table format or with a bar chart or a similar form of graphical representation. A sample distribution table is presented below, showing the frequency distribution for a variable "age".

Basic way of presenting univariate data is to create a frequency distribution of the individual cases, which is to present the number of cases in the sample that belong to each category of values of the variables. This can be done in a table format or with a bar chart or a similar form of graphical representation。

Table 3.1 Frequency distribution table

| Age range | Number of cases | Percent |
|---|---|---|
| Under 18 | 10 | 5 |
| 18-29 | 50 | 25 |
| 29-445 | 40 | 20 |
| 45-65 | 40 | 20 |
| Over 65 | 60 | 30 |
| Valid cases: 200<br>Missing cases: 0 | | |

As shown in the table 3.1, we can see 5% of cases which is also 10 cases are less than 18 years old. 60 cases which are 30% are over 65 years old.

Expect frequency distribution presenting, univariate analysis can also put forward the central tendency, like the mean or average value, median and mode of variables to help us understand the variables.

The probability of firms' bankruptcy may increase with the increase of independent variables, which are dangerous factors; the probability of firms' bankruptcy decrease with the decrease of independent variables, which are protective factors.

It's convenient to do the univariate logistic in SPSS and judging by getting

coefficient, standard error ...and *p* value

## 3.3 logistic regression model

Logit regression analysis are the multivariate techniques which estimate the probability that an event occurs or not, by predicting a binary dependent outcome from several chosen independent variables. logistic regression is used to refer specifically to the problem in which the dependent variable is binary—that is, the number of available categories is two—while problems with more than two categories are referred to as multinomial logistic regression or polytomous logistic regression, or, if the multiple categories are ordered, as ordinal logistic regression.

Logistic regression is used widely in many fields, including the medical and social sciences and finance. For example, the Trauma and Injury Severity Score (TRISS), which is widely used to predict mortality in injured patients, was originally developed using logistic regression. Many other medical scales used to assess severity of a patient have been developed using logistic regression. Logistic regression may be used to predict whether a patient has a given disease (e.g. leukemia; coronary heart disease), based on observed characteristics of the patient (like age, sex, body mass index, blood tests results, family disease and so on; age, blood cholesterol level, systolic blood pressure, relative weight, blood hemoglobin level, smoking (at different levels), and abnormal electrocardiogram.). Another example might be to predict whether an American voter will vote Democratic or Republican, based on age, income, sex, race, state of residence, votes in previous elections, etc. The technique can also be used in engineering, especially for predicting the probability of failure of a given process, system or product. It is also used in marketing applications such as prediction of a customer's propensity to purchase a product or halt a subscription, etc. In economics it can be used to predict the likelihood of a person's choosing to be in the labor force, and a business application would be to predict the probability of a homeowner defaulting on a mortgage or bankruptcy of a company. Conditional

random fields, an extension of logistic regression to sequential data, are used in natural language processing.

Logistic regression can be binomial or multinomial. Binomial or binary logistic regression deals with situations in which the observed outcome for a dependent variable can have only two possible types(again married or unmarried). Multinomial logistic regression deals with situations where the outcome can have three or more possible types (like "disease A" vs. "disease B" vs. "disease C"). In binary logistic regression, the outcome is usually coded as "0" or "1", as this leads to the most straightforward interpretation. If a particular observed outcome for the dependent variable is the obviously possible outcome (referred to as a "success" or a "case" occurs) it is usually coded as "1" and the contrary outcome (referred to as a "failure" or a non-occur) as "0". Logistic regression is used to predict the odds of being a case based on the values of the independent variables (predictors). The odds are defined as the probability that a particular outcome is a case divided by the probability that it is not occur.

If a particular observed outcome for the dependent variable is the obviously possible outcome (referred to as a "success" or a "case" occurs) it is usually coded as "1" and the contrary outcome (referred to as a "failure" or a non-occur) as "0". Logistic regression is used to predict the odds of being a case based on the values of the independent variables (predictors). The odds are defined as the probability that a particular outcome is a case divided by the probability that it is not occur.

A success logistic regression model can be fitted to the predictors using linear regression analysis. The predicted value of the logit will be converted into predicted odds via the inverse of the natural logarithm, which is the exponential function. Thus, although the observed dependent variable in logistic regression is a zero-or-one variable, the logistic regression estimates the odds, as a continuous variable, that the dependent variable is a “success” (occur). In some applications the odds are all that is needed. In others, a specific yes-or-no prediction is needed for whether the dependent variable is or is not a case; this categorical prediction can be based on the computed odds of a success, with predicted odds above some chosen cutoff value being

translated into a prediction of a success.

Here, we take $y_i$ equals to 1 as the company bankrupt (with the probability $P_i$) and $y_i$ equals to 0 as the company not go bankrupt (With the probability $1$-$P_i$).

In logistic regression, we model the probability $P$ that the company go bankrupt by specifying the following

Logistical regression model is standard logistic distribution of error. In logit model, the linear relationship is estimated through an exponential transformation, called: logistic transformation

$$P_i = f(w_i) = 1/(1 + e^{-w_i}) \tag{3.21}$$

Where $w_i$ is independent variables estimated through linear function of financial Indicators $x_{ij}$ (2.18) can be expressed as:

$$w_i = \propto + \sum_{j=1}^{m} \beta_j x_{ij} \tag{3.22}$$

Because of above procedure, the logistical regression model can be expressed by

$$P_i = 1/(1 + e^{-\propto - \Sigma_j \beta_j x_j}) \tag{3.23}$$

And we need to pick up the appropriate $x_i$ to calculate the probability of firm's bankruptcy by formula (3.22).

### 3.3.1 The Wald test (stepwise forward method)

After basic exclude independent variables by univariate analysis, we can put these significant variables into Wald test.

Stepwise procedure of selection of variables is based on measure the statistical significance of the coefficient for the variable. In logistic regression the errors are assumed to follow a binomial distribution, and significance is assessed via likelihood ratio chi-square test.

The selection of variables from a model based on stepwise procedure is to check for the importance of variables, and the importance of a variable is defined in terms of a measure of the statistical significance of the coefficient for the variable. Now,

almost all the major software have the option for stepwise logistic regression which is quite simple.

**Maximum Likelihood Estimation**

The regression coefficients are usually estimated by using maximum likelihood estimation. It is not possible to find a closed-form expression for the coefficient values that maximize the likelihood function unlike linear regression, so that an iterative process must be used instead; for example Newton's method. This process begins with a test to current project and change or modifies it until it is improved enough.

Assume that we have a sample of $n$ independent observations ($x_i$, $y_i$), $i$=1,2,…,n. The method of estimation used in multivariable logistic regression model will be the same as in the univariate situation-maximum likelihood.

- The coefficients could be theoretically estimated by splitting the sample into pools with similar characteristics and by OLS
- …or by maximizing the total likelihood or log-likelihood

If Y is coded as 0 or 1, we have $P(Y = 1|x)$ stands for probability that Y is equal to 1 with given $x$. similarly $P(Y = 0\,|x)$ refers to probability that Y is equal to 0 with given $x$. The contribution to the likelihood function is$1 - \pi(x_i)$, where the quantity $\pi(x_i)$ denotes the value of $\pi(x)$ calculated at $x_i$. Then we have the expression:

$$\pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i} \tag{3.23}$$

And based on expression (3.23), the likelihood function is obtained as the product of the terms like:

$$l(\beta) = \prod_{i=1}^{n} \pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i} \tag{3.24}$$

$$L(\beta) = \ln l(\beta) = \sum_{i=1}^{N} Y_i * \ln\big(\pi(x_i)\big) + (1 - Y_i) * \ln(1 - \pi(x_i)) \tag{3.25}$$

### 3.3.2 Odds ratio

The odds ratio is a measure of association which has found wide use, as it approximates how much more likely or unlikely for the outcome to be present among those with *x*=1 than among those with *x*=0. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.

Odds ratios are used to compare the relative odds of the occurrence of the outcome of interest (like bankruptcy of a firm), given exposure to the variable of interest (like heavy debt history or bad reputation of a company). The odds ratio can also be used to determine whether a particular exposure is a risk factor for a particular outcome, and to compare the degree of various risk factors for that outcome.

Let's say, if y represents the probability of a firm go bankruptcy (bankruptcy or non-bankruptcy) and if x denotes the a smart CEO of the company, then *OR*=0.5 estimates that the firm's bankruptcy is one half as likely to occur with the smart CEO exists and if *OR*=2 means that the firm's bankruptcy is twice as likely to occur with given *x* value.

Table 3.2: values of logistic regression model when independent variable is dichotomous.

| Outcome Variable(Y) | Independent variable (X) | |
|---|---|---|
| | x=1 | x=0 |
| y=1 | $\pi(1) = e^{\beta_0+\beta_1}/(1+e^{\beta_0+\beta_1})$ | $\pi(0) = \frac{e^{\beta_0}}{(1+e^{\beta_0})}$ |
| y=0 | $1-\pi(1) = 1/(1+e^{\beta_0+\beta_1})$ | $1-\pi(0) = 1/(1+e^{\beta_0})$ |
| Total | 1.0 | 1.0 |

$$\mathrm{OR} = \frac{(\frac{e^{\beta_0+\beta_1}}{1+e^{\beta_0+\beta_1}})/(\frac{1}{1+e^{\beta_0+\beta_1}})}{(\frac{e^{\beta_0}}{1+e^{\beta_0}}) \ / \ (\frac{1}{1+e^{\beta_0}})} = e^{\beta_1} \tag{3.26}$$

For a dichotomous variable (have two values) the parameter of interest is the odds ratio. An estimate of this parameter may be obtained from the estimated logistic regression coefficient, regardless of how the variable is coded.

And about confidence intervals, usually the 95% confidence interval (CI) is used to estimate the precision of the OR. A large CI indicates a low level of precision of the OR, whereas a small CI indicates a higher precision of the OR. It is important to note however, that unlike the *p* value, the 95% CI does not report a measure's statistical significance. In practice, the 95% CI is often used as a proxy for the presence of statistical significance if it does not overlap the null value (e.g. OR=1).

### 3.3.3 Classification table

For binary response data, the response is either an event occurs or not occurs. In binary logistic, the response with ordered Value 1 is regarded as the event occurs, and the response with ordered Value 0 is the event not occurs. Logistic regression models the probability of the event. From the fitted model, a predicted event probability can be computed for each observation.

To obtain the dichotomous variable we must define a cut point, and compare each estimated probability to the cut point. If the estimated probability is bigger than the cut point then we let the derived variable be equal to 1; vice versa, if the estimated probability is smaller than the cut point, the derived variable will be defined as 0.

A method to compute a reduced-bias estimate of the predicted probability is given in the section predicted probability of an event for classification. If the predicted event probability exceeds or equals some cut point value, the observation is predicted to be an event observation; otherwise, it is predicted as a nonevent. A frequency table can be obtained by cross-classifying the observed and predicted

responses.

The accuracy of the classification is measured by its sensitivity (the ability to predict an event occurs correctly) and specificity (the ability to predict a non-occur event correctly). Sensitivity is the proportion of event responses that were predicted to occur. Specificity is the proportion of nonevent responses that were predicted to be not occurs. Logistic regression also computes three other conditional probabilities: false positive rate, false negative rate, and rate of correct classification. The false positive rate is the proportion of predicted event responses that were observed as non-occur. The false negative rate is the proportion of predicted non-occur responses that were observed as event occurs. An example of classification table is below shown in Table3.3.

Table 3.3: Classification table

$\pi$=0.55

| Classified | Observed | | Total |
|---|---|---|---|
| | DFREE=1<br>Drug Free | DFREE=0<br>Returned to Drug Use | |
| DFREE=1 | 15 | 12 | 27 |
| DFREE=0 | 301 | 247 | 548 |
| Total | 316 | 259 | 575 |

Sensitivity=15/316=1.7%; Specificity=247/259=95.4%

It's simply to calculate the prior probability of default, the false positive rate and false negative rate:

$$P_{F+} = P_r(-\mathrm{B}|\mathrm{A}) = \frac{P_r(\mathrm{A}|-\mathrm{B})[1-P_r(B)]}{P_r(\mathrm{A}|-\mathrm{B})+P_r(\mathrm{B})[P_r(\mathrm{A}|\mathrm{B})-P_r(\mathrm{A}|-\mathrm{B})} \tag{3.27}$$

$$P_{F-} = P_r(\mathrm{B}|-\mathrm{A}) = \frac{[1-P_r(\mathrm{A}|\mathrm{B})]P_r(B)}{1-P_r(\mathrm{A}|-\mathrm{B})-P_r(\mathrm{B})[P_r(\mathrm{A}|\mathrm{B})-P_r(\mathrm{A}|-\mathrm{B})]} \tag{3.28}$$

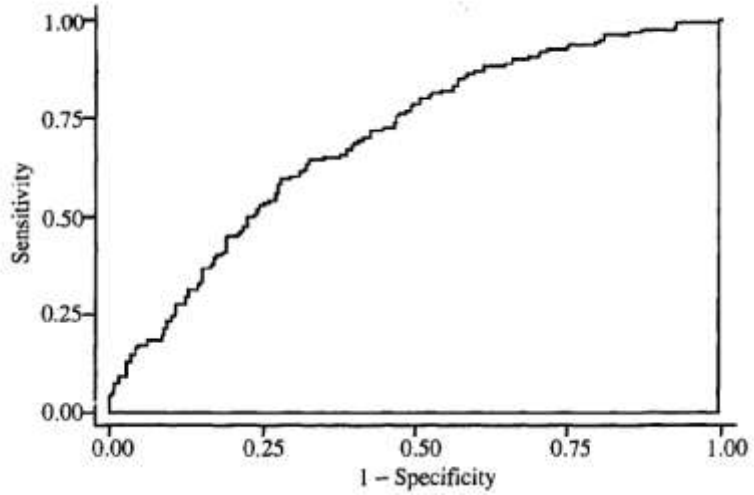
### 3.3.4 ROC curve

Receiver operating characteristic curve (ROC curve) is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate against the false positive rate at various threshold settings which is sensitivity against specificity.

A ROC space is defined by FPR and TPR as *x* and *y* axes respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs). Since TPR is equivalent to sensitivity and FPR is equal to 1 − specificity, the ROC graph is sometimes called the sensitivity vs (1 − specificity) plot. Each prediction result or instance of a confusion matrix represents one point in the ROC space. Combine with the prediction and performances to generate the ROC curve under evaluate method.

The area under the ROC which is also called AUC, when using normalized units, the area under the curve is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one (assuming 'positive' ranks higher than 'negative'). This can be seen as follows: the area under the curve is given by. The area of actual model is shown in Figure 3.1.

Figure 3.1 ROC curve[4] (Plot of sensitivity versus 1-specificity for all possible cut points.




[4] See in Bibliography [4], P163

AUC=0.5, It suggests no discrimination.

0.5＜AUC＜0.7, which means the model has a few veracity.

0.7＜AUC＜0.9, which means the model has a certain extent veracity, excellent.

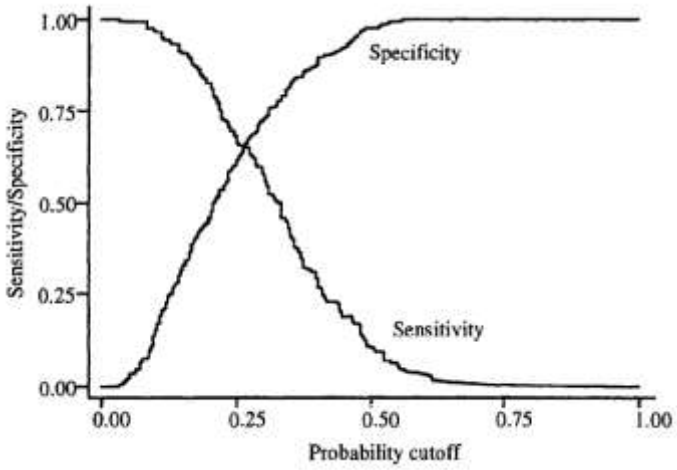
The more closely to 1, the better the accuracy is the model.

We can also use ROC curve to judge the better method by compare the area under the ROC curve, the bigger the better.

The test measurements may contain missing values and two methods are provide d to handle missing values when comparing ROC areas – pairwise deletion and casew ise deletion. This is described in detail later.

**Sensitivity and specificity versus criterion value**

Figure 3.2: Sensitivity and specificity versus criterion value[5]




When you select a higher criterion value, the false positive fraction will decrease with increased specificity but on the other hand the true positive fraction and sensitivity will decrease:

[5] Also see in Bibliography [4]

# 4. Empirical Scoring Model

In this section, by using method described in chapter three empirical model will be made after filter the significant variables which are financial ratios.

## 4.1 Input data description

Data description process is mainly to do the basic data analysis, to describe basic characteristic of variables. We can use SPSS to generate the descriptive statistics like mean, std. deviation, minimum, maximum and variance. Though this statistics, we can get comprehensive understanding of the feature of variables.

Table 4.1: Data descriptive Statistics

| | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
|---|---|---|---|---|---|---|
| ROE | 372 | -41.50 | 40.50 | .4120 | 3.82339 | 14.618 |
| ROA (EBITDA) | 371 | -1067.00 | 10.95 | -3.5947 | 55.96911 | 3132.542 |
| ROA | 371 | -1195.00 | 10.95 | -3.9937 | 62.57730 | 3915.918 |
| ROS | 369 | -120.25 | 537.12 | .7020 | 28.92417 | 836.608 |
| D/A | 371 | -19.72 | 1200.00 | 6.5487 | 67.37519 | 4539.417 |
| Bank loans/A | 371 | -.01 | 144.20 | .5765 | 7.49637 | 56.196 |
| DSI | 369 | .00 | 1992.59 | 51.9717 | 181.38231 | 32899.543 |
| DSO | 369 | -367.20 | 1239750 | 3784.9 | 64587.52 | 4171548358 |
| A*360/Sales | 369 | -56.90 | 8404200 | 26944.0 | 438795.87 | 192541819336 |
| current ratio | 372 | -10.81 | 214.00 | 4.8711 | 17.98192 | 323.349 |
| Quick ratio | 372 | -10.10 | 214.00 | 4.5060 | 17.97285 | 323.023 |
| Cash ratio | 372 | -8.81 | 214.00 | 3.1954 | 17.04733 | 290.611 |
| working capital to A | 371 | -1199.00 | 50.90 | -5.3330 | 68.55948 | 4700.403 |

| | | | | | | |
|---|---|---|---|---|---|---|
| gross profit margin | 369 | -71.00 | 1.15 | -.3839 | 5.41579 | 29.331 |
| Financial leverage | 372 | -98.17 | 934.95 | 7.8476 | 61.19683 | 3745.051 |
| interests/ EBIT | 372 | -17.13 | 5.79 | -.0369 | 1.17444 | 1.379 |
| capital employed/FA | 297 | -1413.14 | 1001.98 | .2641 | 115.51873 | 13344.576 |
| current liabilities/A | 371 | -.28 | 1200.00 | 5.6830 | 66.77755 | 4459.241 |
| D/E | 372 | -82.99 | 933.95 | 6.7966 | 60.94417 | 3714.191 |
| Bank loans/E | 372 | -80.30 | 303.09 | .6147 | 17.46079 | 304.879 |
| Valid N (listwise) | 296 | | | | | |

According to Table 4.1, first we can see the number of each variable. The number of "capital employed/FA" is 297 which are far more less than the number of other variables. And its maximum and minimum value shows the extreme situation of data, the value of std. deviation and variance also proved it. This kind of variable is not a good one for deeper analysis.

Arithmetic Mean value is one of most used measurement in central tendency. Variance tests the dispersion which is increase of each unit of its arithmetic mean of the correlations. A small variance indicates that the data points tend to be very close to the mean (expected value) and hence to each other, while a high variance indicates that the data points are very spread out around the mean and from each other. Standard deviation is the square root of a variance. Like variables "current ratio", "quick ratio" and "cash ratio" have got similar mean value, std. deviation value, we have no need to include all three variables in analysis.

## 4.2 Univariable analysis

Before the Wald test, we can do a univariable analysis to delete some insignificant financial indicators or multiple linear problems.

Table 4.2: Univariables analysis table

| | | B | S.E. | Wald | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Lower | Upper |
| x1 | ROE | 0.022 | 0.029 | 0.58 | 0.446 | 1.022 | 0.966 | 1.082 |
| x2 | ROA(EBITDA) | -1.062 | 0.288 | 13.597 | 0 | 0.346 | 0.197 | 0.608 |
| x3 | ROA | -1.009 | 0.277 | 13.318 | 0 | 0.364 | 0.212 | 0.627 |
| x4 | ROS | -0.507 | 0.199 | 6.51 | 0.011 | 0.602 | 0.408 | 0.889 |
| x5 | D/A | 1.051 | 0.195 | 29.12 | 0 | 2.862 | 1.953 | 4.193 |
| x6 | Bankloans/A | 1.985 | 0.515 | 14.871 | 0 | 7.277 | 2.654 | 19.953 |
| x7 | DSI | 0.002 | 0.001 | 3.559 | 0.059 | 1.002 | 1 | 1.004 |
| x8 | DSO | 0 | 0 | 0.422 | 0.516 | 1 | 1 | 1 |
| x9 | A*360/Sales | 0 | 0 | 0.445 | 0.505 | 1 | 1 | 1 |
| x10 | current ratio | -0.414 | 0.085 | 23.815 | 0 | 0.661 | 0.56 | 0.781 |
| x11 | Quick ratio | -0.408 | 0.088 | 21.633 | 0 | 0.665 | 0.56 | 0.79 |
| x12 | Cash ratio | -1.01 | 0.201 | 25.36 | 0 | 0.364 | 0.246 | 0.54 |
| x13 | Working capital/CA | -0.223 | 0.078 | 8.216 | 0.004 | 0.8 | 0.687 | 0.932 |
| x14 | gross profit margin | -0.027 | 0.025 | 1.165 | 0.28 | 0.973 | 0.927 | 1.022 |
| x15 | Financial leverage | 0 | 0.002 | 0.005 | 0.946 | 1 | 0.997 | 1.003 |
| x16 | interests/EBIT | -0.181 | 0.131 | 1.914 | 0.166 | 0.835 | 0.646 | 1.078 |
| x17 | capital employed/FA | -0.021 | 0.009 | 5.679 | 0.017 | 0.98 | 0.963 | 0.996 |
| x18 | current liabilities/A | 2.012 | 0.289 | 48.548 | 0 | 7.477 | 4.246 | 13.168 |
| x19 | D/E | 0 | 0.002 | 0.004 | 0.951 | 1 | 0.997 | 1.003 |
| x20 | Banklolans/E | 0.005 | 0.007 | 0.452 | 0.502 | 1.005 | 0.991 | 1.018 |

Based on Table 4.2, we can see the red part *p* value are all above 0.05, which means variables x1, x7,x8,x9,x14,x15,x16,x19 and x20 are not significant, we can delete them. The rest can be continuing test in the Wald test.

And *B* is the coefficient of variables to prediction of firms' default; here coefficients which are negative mean variables have a negative effect on *p* value, we called them protection factor; vice versa, the coefficients which are positive denotes a positive effect on *p* value and we call the risk factors.

## 4.3 Multivariable regression analysis

In this section, we will build the estimate logistic model by theory describe in chapter three on SPSS.

### 4.3.1 The Wald test

First we use the Wald test to choose independent variable into the model step by step and then pick up the best fitting model by the regression analysis result.

Table 4.3: Encoding information of dependent variables

| Dependent Variable Encoding | |
|---|---|
| Original Value | Internal Value |
| pro. of non-bankruptcy | 0 |
| pro. of bankruptcy | 1 |

Initial value 0 indicates that the company is not bankruptcy, and value 1 denotes the firm has probability of bankruptcy.

Table 4.4: Test results of no independent variable but constant.

| Variables in the Equation | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | B | S.E. | Wald | df | Sig. | Exp(B) |
| Step 0 | Constant | -.176 | .117 | 2.278 | 1 | .131 | .839 |

We can see in the Table 4.4, the *B* equals to logit value of the firm's bankruptcy, *S.E.* is the standard error. Wald chi-square value is 2.278 and sig. is the *p* value of Wald chi-square at 1 degree of free.

Table 4.5: Omnibus Tests of Model Coefficients

| | Chi-square | df | Sig. |
|---|---|---|---|
| Step | 5.653 | 1 | .017 |

| | | | |
|---|---|---|---|
| Block | 119.444 | 4 | .000 |
| Model | 119.444 | 4 | .000 |

We can see the significant values are all smaller than 0.05, which means the process is ok. Let's continue.

Table 4.6: Model summary

| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|---|---|---|---|
| 1 | 308.414a | .286 | .382 |
| 2 | 301.100a | .303 | .405 |
| 3 | 294.265a | .319 | .427 |
| 4 | 288.612a | .332 | .444 |

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

As shown in Table 4.6, the -2 log likelihood value of final model is the lowest 288.6 and two $R^2$ is higher which means the model has better goodness of fit.

Table 4.7:Hosmer and Lemeshow Test

| Chi-square | df | Sig. |
|---|---|---|
| 11.778 | 8 | .161 |

The sig. value is higher than 0.05, which means the model is significant.

| | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Lower | Upper |
| ROAEBITDA | -1.351 | .555 | 5.923 | 1 | .015 | .259 | .087 | .769 |
| DA | 1.005 | .481 | 4.365 | 1 | .037 | 2.732 | 1.064 | 7.015 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| BankloansA | 2.083 | .814 | 6.543 | 1 | .011 | 8.028 | 1.627 | 39.602 |
| currentliabilitiestoA | 1.935 | .552 | 12.311 | 1 | .000 | 6.925 | 2.349 | 20.412 |
| Constant | -2.363 | .349 | 45.751 | 1 | .000 | .094 | | |

Table 4.8: Logistic regression analysis result

| | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Lower | Upper |
| ROA(EBITDA) | -1.351 | .555 | 5.923 | 1 | .015 | .259 | .087 | .769 |
| DA | 1.005 | .481 | 4.365 | 1 | .037 | 2.732 | 1.064 | 7.015 |
| Bank loans/A | 2.083 | .814 | 6.543 | 1 | .011 | 8.028 | 1.627 | 39.602 |
| CL/A | 1.935 | .552 | 12.311 | 1 | .000 | 6.925 | 2.349 | 20.412 |
| Constant | -2.363 | .349 | 45.751 | 1 | .000 | .094 | | |

Log likelihood=-144.306 R square=0.33

Then we can write the prediction model based on the analysis result:

$$y = -1.351x_2 + 1.005x_5 + 2.083x_6 + 1.935x_{19} - 2.363 \quad (4.1)$$

Logit model:

$$p_i = \frac{1}{1+\exp[-(-1.351x_2+1.005x_5+2.083x_6+1.935x_{19}-2.363)}} \quad (4.2)$$

Here x2 refers to return on assets, x5 means debt to assets ratio, x6 is bank loan to assets ratio and x19 is current liability to assets ratio. These four independent variables have important effect on firms' bankruptcy, the probability of bankruptcy increases as the increase of variables x2, and decreases as the increase of variables x5, x6 and x19.

## 4.3.2 Odds ratio

Now, we may find out if our model is accuracy or not. First we calculated the odds ratio of each independent variables and build up the estimated odds ratio and 95 percent confidence limits based on the model.

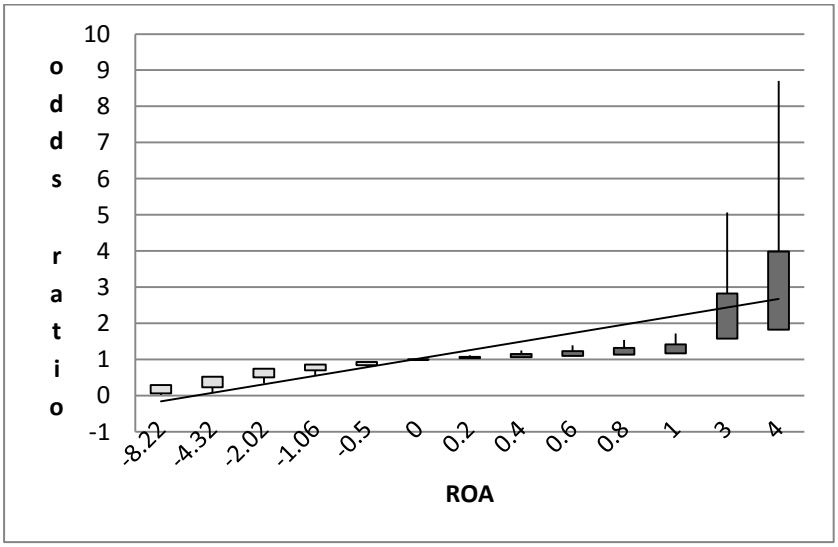
Table 4.9: Odds ratio of ROA by Stata

After steps on Stata, we can get the odds ratio and the standard error of the variable.

| Logistic regression | | | | | | |
|---|---|---|---|---|---|---|
| Proablility of Bankruptcy | Odds Ratio | Std.Err. | z | P>\|z\| | [95% Conf. Interval] | |
| ROA | 0.345656 | 0.099581 | -3.69 | 0 | 0.1965257 | 0.607953 |
| _cons | 0.805111 | 0.087563 | -1.99 | 0.046 | 0.6505484 | 0.996396 |

| Proablility of Bankruptcy | Odds Ratio | Std.Err. | z | P>\|z\| | [95% Conf. Interval] | |
|---|---|---|---|---|---|---|
| Bankloans/A | 7.276601 | 3.744949 | 3.86 | 0 | 2.653703 | 19.95285 |
| _cons | 0.638326 | 0.078856 | -3.63 | 0 | 0.5010586 | 0.813199 |
| Logistic regression | | | | | | |

Figure 4.1 Estimated odds ratio and 95 percent confidence limits for return on assets based on the model



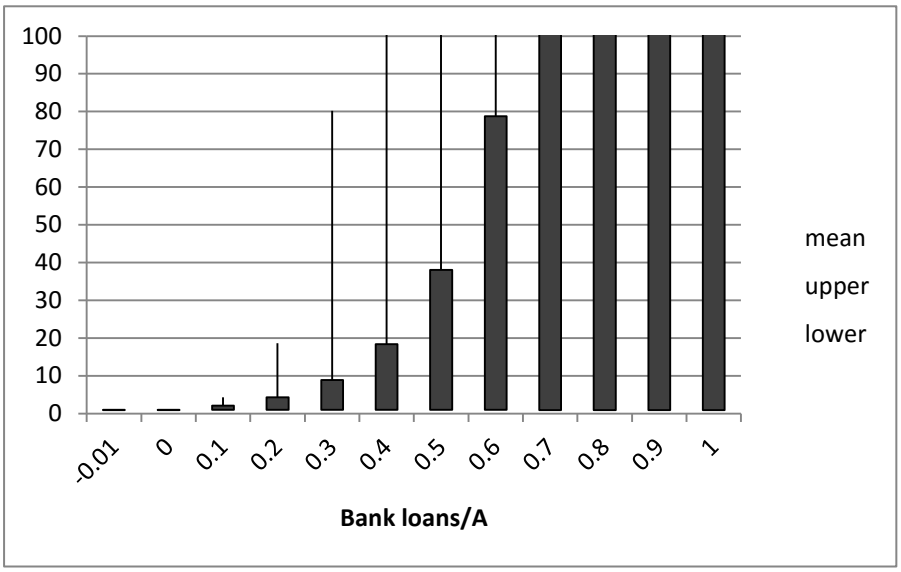
See on Chart 4.1, before ROA equals to 0, odds ratio are all smaller than 1, when ROA is above 0, we can see the odds ratio is higher than 1 which means the probability of bankruptcy is more likely to occur.

Table 4.10: odds ratio of Bank loan/A by Stata

Again we can do simply calculation of bank loan to assets on Stata.

| Logistic regression | | | | | | |
|---|---|---|---|---|---|---|
| Probability of Bankruptcy | Odds Ratio | Std.Err. | z | P＞\|z\| | [95% Conf. Interval] | |
| D/A | 1.324017 | 0.21503 | 6.16 | 0 | 0.9025659 | 1.745467 |
| _cons | -1.66425 | 0.224473 | -7.41 | 0 | -2.104208 | -1.22429 |

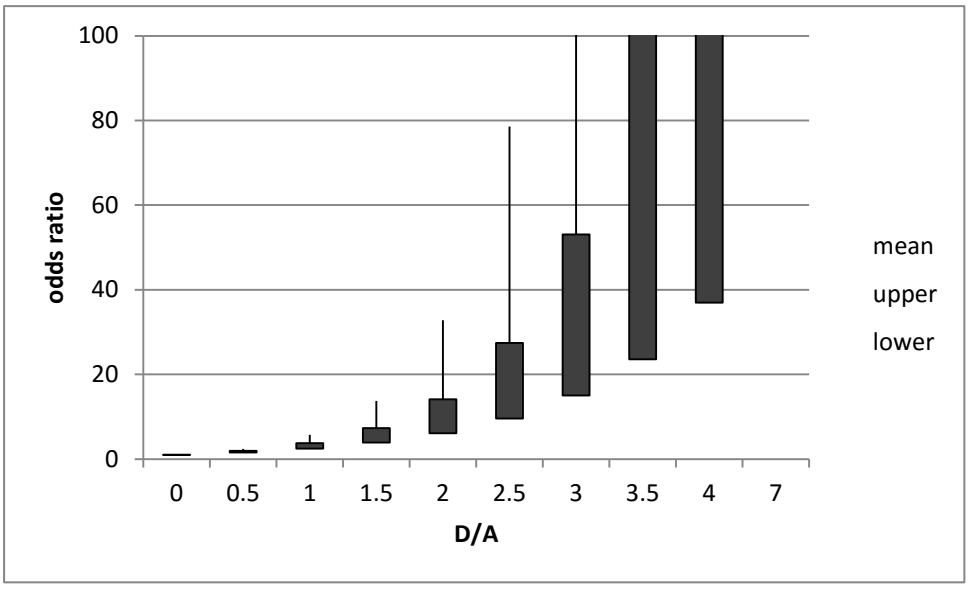
Figure 4.2: Estimated odds ratio and 95 percent confidence limits for bank loan to assets base on the model




Here we see that after bank loan /A is above 0.1, the lower odds ratio is above 1, which means the probability of bankruptcy is more likely to occur.

Again we get the odds ratio of D/A 1.324 and standard error 0.215 from STATA in the same way. And results show in Figure 4.3.

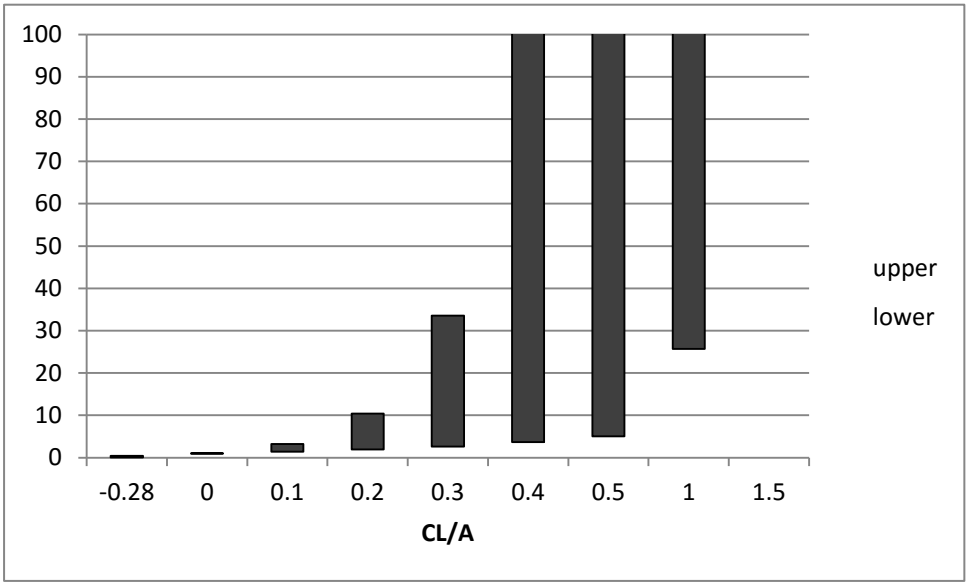
Figure 4.3 Estimated odds ratio and 95 percent confidence limits for debt to assets base on the model




The estimated odds ratio is all higher than 1 which indicates a more likely probability of firm's bankruptcy

The odds ratio of CL/A is 7.477 and standard error is 2.159 according to Table 4.12.

Figure 4.4 Estimated odds ratio and 95 percent confidence limits for current liability to assets base on the model




When independent variable equals to -0.28, the mean odds ratio is 0.123 which is smaller than 1, the probability of company's bankruptcy is less likely to happen. Vice

versa, after odds ratio is above 1 which indicates the more likely occur of firm's bankruptcy.

### 4.3.3 Classification table

Table 4.13: classification table

| Observed | | Predicted | | | |
|---|---|---|---|---|---|
| | | probability of b | | | Percentage Correct |
| | | 0（D） | 1（~D） | total | |
| probability of b | 0（+） | 137 | 24 | 161 | 85.1 |
| | 1（-） | 43 | 92 | 135 | 68.1 |
| | | 180 | 116 | | 77.4 |

The cut value is 0.500

According to first Table 4.13, we can easily figure it out that there are 161 non-bankruptcy companies and 135 bankruptcy companies taken into account in our model. 137 of non-bankruptcy are successful predict by our model, so do 92 bankruptcy firms, which is again 24 non-bankruptcy companies are predict false as well as 43 bankruptcy firms. Percentage correct is 77.4 percent, which is a good one.

Table 4.14: calculations due to classification table

| | | |
|---|---|---|
| sensitivity | Pr( +\| D) | 76.11% |
| specificity | Pr( -\|~D) | 79.31% |
| positive predictive value | Pr( -\|~D) | 85.09% |
| negative predictive value | Pr(~D\| -) | 68.15% |
| False + rate for true ~D | Pr( +\|~D) | 20.69% |
| False - rate for true D | Pr( -\|~D) | 23.89% |
| False + rate for classified + | Pr(~D\| +) | 14.91% |
| False - rate for classified - | Pr(D\| -) | 31.85% |

More specifically, we calculate the sensitivity and specificity as well as positive

predictive value proportion, negative predictive value based on formula described in chapter three. Sensitivity of non-bankruptcy is 137/180=76.1% and specificity =92/116=79.3%.

For example, false + rate for true ~D indicates that the wrong predict non-bankruptcy firms to total predicted bankruptcy firms. False - rate for true D refers to the wrong predict bankruptcy firms to total predicted non-bankruptcy firms.

### 4.3.4 ROC curve

We have 1-speciality as abscissa axis and sensitivity as vertical coordinates to generate the ROC curve.

Table 4.15 :Area Under the Curve

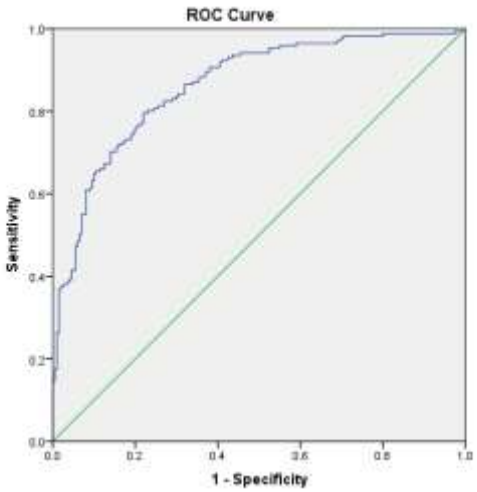
Test Result Variable(s): Predicted probability

| Area | Std. Errora | Asymptotic Sig.b | Asymptotic 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| .863 | .019 | .000 | .825 | .900 |

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Chart 4.5: ROC curve




ROC area: 0.863

X-axis represents specificity, which is the relation of default incorrect classified with total default group (the incorrect predicted percentage of probability of default firms), and y-axis informs to sensitivity, it indicates the correct predicted percentage of probability of non-default firms. The more ROC curve is closer to the upper left corner of the figure, the more effective of our model is to estimate predicted probability of default, then we can also measure if our model is significant or not.

According to Table 4.15, the area under ROC curve is 0.863, 95% confidence interval is 0.825 to 0.900, which is very close to the ideal type--1, standard error is 0.019 and sig. value is 0.00, our model has a good accuracy on estimate the probability of firms' bankruptcy. The selected independent variables are significant to our model.

# 5. Conclusion

In this thesis, we are trying to build a scoring model by using logistic regression method to classify a company is a good or a bad one. After analysis nearly 400 companies and pick up the most relatively or significance variables to build up the model.

At first, we have done the univariable analysis, to delete some less relatively or important independent variables like gross profit margin and financial leverage. Then we put the variables into the Wald test step by step to select the most important variables to generate the model.

After model generated, we have classification table and ROC curve to evaluate if our model can explain the data and the prediction of firms' bankruptcy. Fortunately, we have got 77.4 percentage predictions correct in classification table and 0.863 in ROC curve, which means our model can explain the topic to some extent and help us make prediction about companies, avoid risk and to be a smart investors.

Although we study in credit scoring model of analysis financial status of firms, but the credit risk management is a huge work and changes all the time by various factors, as we have limited time and ability to use more methods, our model of prediction could be more significant and powerful.

# Bibliography

[1] John Hull, Izzy Nelken, and Alan White: Merton's Model, Credit Risk, and Volatility Skews Sep.2004

[2] CRAMER, Jan. Logit models from economics and other fields. Cambridge: Cambridge University Press, 2003, 173 p. ISBN 0-521-81588-6.

[3] GOURIEROUX, Christian and Joann JASIAK. Econometrics of individual risk: credit, insurance and marketing. Princenton: Princeton University Press, 2007, 241 p. ISBN 978-0-691-12066-9.

[4] HOSMER, David W. and Stanley LEMESHOW. Applied logistic regression. 2nd ed. New York: Wiley, 2000, 375 p. ISBN 0-471-35632-8.

[5] H. Bühlmann A. Pelsser W. Schachermayer H. Waters D. Filipovic, Chair. Concentration Risk in Credit Portfolios springer 2009 ISBN 978-3-540-70869-8

[6] Aijun Zhang. Statistical Methods in Credit Risk Modeling 2009 The University of Michigan

[7] Jose A. Lopez. Evaluating Credit Risk Models 1999 Research and Market Analysis Group Federal Reserve Bank of New York

[8] Edward I. Altman* PREDICTING FINANCIAL DISTRESS OF COMPANIES: REVISITING THE Z-SCORE AND ZETA 2000.7

[9] FRIDSON, Martin and ALVAREZ, Fernando. Financial Statement Analysis. 4th ed. New York: Wiley Finance, 2011.193p. ISBN: 978-0470640036

[10] TIROLE, Jean. The Theory of Corporate Finance.1st ed. New York: Princeton University Press, 2005.640p. ISBN: 978-0691125565

[11] DLUHOSOVA, Dana, TICHY, Tomas and ZMESKAL, Zdenek. Financial Models. 1st ed.VSB-Technical University of Ostrava, 2004. 254p. ISBN: 80-24807548

[12] Warren Miller Morningstar, Inc., Comparing Models of Corporate Bankruptcy Prediction: Distance to Default vs. Z-Score, 2009.7

[13] Almeida, H., and Philippon, T. "The Risk Adjusted Cost of Financial Distress." Journal of Finance, 62 (2007), pp. 2557-2586.

[14] Begley, J., Ming, J., and Watts, S. "Bankruptcy Classification Errors in the 1980s: An Empirical Analysis of Altman's and Ohlson's Models." Review of Accounting Studies, 1 (1996), pp. 267-284.

[15] Hotchkiss, E., John, K., Mooradian, R., and Thorburn, K., Elsevier, B.V., "Bankruptcy and the Resolution of Financial Distress." In Handbook of Empirical Corporate Finance (2008).

[16] Shumway, T. "Forecasting Bankruptcy More Accurately: A Simple Hazard Model." Journal of Business, 74 (2001), pp. 101-124

[17] Ohlson, J. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." Journal of Accounting Research, 18 (1980), pp. 109-131.

[18] Hillegeist, S., Keating, E., Cram, D., and Lundstedt, K. “Assessing the Probability of Bankruptcy.” Review of Accounting Studies, 9 (2004), pp. 5-34

[19] Kealhofer, S., and Kurbat, M. “The Default Prediction Power of the Merton Approach: Relative to Debt Ratings and Accounting Variables.” KMV LLC, 2001.

[20] Duffie, D., Saita, L., and Wang, K. “Multi-Period Corporate Failure Prediction with Stochastic Covariates.” Journal of Financial Economics, 83 (2007), pp. 635-665.

## Electronic reference

[1] U.S Securities and Exchange Commission.
http://www.sec.gov/index.htm

[2] Logistic regression analysis.
http://blog.sina.com.cn/s/blog_51c4baac0100w125.html

[3] Bond Credit Rating Table.
http://learnbonds.com/bond-credit-ratings-table/6891/

[4] SAS data analysis examples.
http://www.ats.ucla.edu/stat/sas/dae/logit.htm

[5] Residual analysis
http://blog.renren.com/share/227827897/10086451002

[6] SPSS logistic regression.
http://jingyan.baidu.com/article/fdffd1f81f1c0ff3e98ca11e.html?qq-pf-to=pcqq.c2c

[7] Predicting financial distress of companies.
http://pages.stern.nyu.edu/~ealtman/Zscores.pdf

[8] Odds ratio.
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/

[9] Comparing Models of Corporate Bankruptcy Prediction: Distance to Default vs. Z-Score
http://corporate.morningstar.com/us/documents/MethodologyDocuments/MethodologyPapers/CompareModelsCorporateBankruptcyPrediction.pdf

[10] Modelling of prediction distress of firms.
http://www.ckyk.cn/upload_files/20140307/2316095.pdf

[11] ROC curve analysis in MedCalc.
https://www.medcalc.org/manual/roc-curves.php

[12] Expected Loss & Unexpected Loss
http://riskarticles.com/credit-risk-how-to-calculate-expected-loss-unexpected-loss/

# List of Abbreviations

ROE: Return on equity

ROA: Return on assets

ROS: Return on sales

CL: Current liability

AUC: Area under ROC curve

DSO: Day's sales outstanding in accounts receivables

DSI: Day's sale of inventory

MDA: Mechanics Dynamics Aesthetics

EBIT: Earnings before interest and taxes

S: Sales

A: Assets

D: Debt

OR: Odds ratio

LDA: Linear discriminant analysis

# Declaration of Utilization of Results from a Diploma Thesis

Herewith I declare that

- I am informed that Act No. 121/2000 Coll. – the Copyright Act, in particular, Section 35 – Utilization of the Work as a Part of Civil and Religious Ceremonies, as a Part of School Performances and the Utilization of a School Work – and Section 60 – School Work, fully applies to my diploma (bachelor) thesis;
- I take account of the VSB – Technical University of Ostrava (hereinafter as VSB-TUO) having the right to utilize the diploma (bachelor) thesis (under Section 35(3)) unprofitably and for own use ;
- I agree that the diploma (bachelor) thesis shall be archived in the electronic form in VSB-TUO's Central Library and one copy shall be kept by the supervisor of the diploma (bachelor) thesis. I agree that the bibliographic information about the diploma (bachelor) thesis shall be published in VSB-TUO's information system;
- It was agreed that, in case of VSB-TUO's interest, I shall enter into a license agreement with VSB-TUO, granting the authorization to utilize the work in the scope of Section 12(4) of the Copyright Act;
- It was agreed that I may utilize my work, the diploma (bachelor) thesis or provide a license to utilize it only with the consent of VSB-TUO, which is entitled, in such a case, to claim an adequate contribution from me to cover the cost expended by VSB-TUO for producing the work (up to its real amount).

Ostrava dated 15.04.2016

Yu Zhang

Yu Zhang

# List of Annexes

Annex 1: Full steps in Wald test

| | | B | S.E. | Wald | df | Sig. | Exp(B) |
|---|---|---|---|---|---|---|---|
| Step 1[a] | DA | 2.730 | .392 | 48.412 | 1 | .000 | 15.329 |
| | Constant | -2.494 | .342 | 53.214 | 1 | .000 | .083 |
| Step 2[b] | currentliabilitieslA | 1.319 | .489 | 7.288 | 1 | .007 | 3.741 |
| | DA | 1.916 | .468 | 16.777 | 1 | .000 | 6.797 |
| | Constant | -2.550 | .343 | 55.182 | 1 | .000 | .078 |
| Step 3[c] | currentliabilitieslA | 1.915 | .534 | 12.873 | 1 | .000 | 6.788 |
| | DA | 1.299 | .494 | 6.907 | 1 | .009 | 3.666 |
| | BankloansA | 2.055 | .803 | 6.549 | 1 | .010 | 7.810 |
| | Constant | -2.654 | .348 | 58.193 | 1 | .000 | .070 |
| Step 4[d] | currentliabilitieslA | 1.935 | .552 | 12.311 | 1 | .000 | 6.925 |
| | ROAEBITDAA | -1.351 | .555 | 5.923 | 1 | .015 | .259 |
| | DA | 1.005 | .481 | 4.365 | 1 | .037 | 2.732 |
| | BankloansA | 2.083 | .814 | 6.543 | 1 | .011 | 8.028 |
| | Constant | -2.363 | .349 | 45.751 | 1 | .000 | .094 |

Annex 2: full steps of classification table

**Classification Table[a]**

| Observed | | | Predicted | | |
|---|---|---|---|---|---|
| | | | proablity of bankcrupcy | | Percentage Correct |
| | | | 0. | pro. of bankcrupcy | |
| Step 1 | proablity of bankcrupcy | 0. | 135 | 26 | 83.9 |
| | | pro. of bankcrupcy | 47 | 88 | 65.2 |
| | Overall Percentage | | | | 75.3 |
| Step 2 | proablity of bankcrupcy | 0. | 131 | 30 | 81.4 |
| | | pro. of bankcrupcy | 49 | 86 | 63.7 |
| | Overall Percentage | | | | 73.3 |
| Step 3 | proablity of bankcrupcy | 0. | 134 | 27 | 83.2 |
| | | pro. of bankcrupcy | 43 | 92 | 68.1 |
| | Overall Percentage | | | | 76.4 |
| Step 4 | proablity of bankcrupcy | 0. | 137 | 24 | 85.1 |

| | | | | |
|---|---|---|---|---|
| | pro. of bankcrupcy | 43 | 92 | 68.1 |
| Overall Percentage | | | | 77.4 |

a. The cut value is .500

## Annex3：Hosmer and Lemeshow Test

| Step | Chi-square | df | Sig. |
|---|---|---|---|
| 1 | 13.915 | 8 | .084 |
| 2 | 18.734 | 8 | .016 |
| 3 | 9.015 | 8 | .341 |
| 4 | 11.778 | 8 | .161 |

## Annex4: log likelihood and R square

| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|---|---|---|---|
| 1 | 308.414[a] | .286 | .382 |
| 2 | 301.100[a] | .303 | .405 |
| 3 | 294.265[a] | .319 | .427 |
| 4 | 288.612[a] | .332 | .444 |

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

## Annex5: Omnibus Tests of Model Coefficients

| | | Chi-square | df | Sig. |
|---|---|---|---|---|
| Step 1 | Step | 99.642 | 1 | .000 |
| | Block | 99.642 | 1 | .000 |
| | Model | 99.642 | 1 | .000 |
| Step 2 | Step | 7.314 | 1 | .007 |
| | Block | 106.956 | 2 | .000 |
| | Model | 106.956 | 2 | .000 |
| Step 3 | Step | 6.836 | 1 | .009 |
| | Block | 113.792 | 3 | .000 |
| | Model | 113.792 | 3 | .000 |
| Step 4 | Step | 5.653 | 1 | .017 |
| | Block | 119.444 | 4 | .000 |

| | | | |
|---|---|---|---|
| Model | 119.444 | 4 | .000 |

Annex6: Table statistics of bankruptcy data

| | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| ROE | 172 | -16.55 | 24.62 | .5783 | 3.13862 |
| ROA (EBITDA/A) | 171 | -1067.00 | 10.95 | -8.0035 | 82.34853 |
| ROA= EBIT/A | 171 | -1195.00 | 10.95 | -8.8219 | 92.08171 |
| ROS = EBIT/S | 170 | -120.25 | 8.89 | -1.6860 | 10.61126 |
| D/A | 171 | -19.72 | 1200.00 | 13.4850 | 98.94354 |
| Bank loans/A | 171 | .00 | 144.20 | 1.1495 | 11.02867 |
| Inventories*360/Sales | 170 | .00 | 1992.59 | 74.6836 | 241.72768 |
| receivables*30/Sales | 170 | -367.20 | 1239750.00 | 7801.3239 | 95103.71301 |
| A*360/Sales | 170 | -56.90 | 8404200.00 | 54086.8501 | 645363.60930 |
| current ratio | 172 | -3.53 | 11.86 | 1.1377 | 1.50510 |
| Quick ratio | 172 | -3.53 | 11.86 | .9055 | 1.46842 |
| Cash ratio | 172 | -5.97 | 8.78 | .1592 | .87207 |
| capital employed/Fixed assets | 135 | -1413.14 | 313.45 | -19.2113 | 138.57038 |
| (current assets-current liabilities)/current assets | 171 | -1199.00 | 50.90 | -11.9332 | 100.73708 |
| gross profit/sales | 170 | -71.00 | 1.15 | -.7512 | 6.21778 |
| Financial leverage | 172 | -98.17 | 408.15 | 8.0806 | 52.35056 |
| interests/ EBIT | 172 | -17.13 | 5.79 | -.1391 | 1.61080 |
| capital employed/Fixed assets | 135 | -1413.14 | 313.45 | -19.2113 | 138.57038 |

| | | | | | |
|---|---|---|---|---|---|
| current liabilities/A | 171 | -.28 | 1200.00 | 11.8495 | 98.15396 |
| D/E | 172 | -82.99 | 407.15 | 7.0037 | 51.70983 |
| Bank loans/E | 172 | -80.30 | 303.09 | 1.3130 | 25.42452 |

Annex7: statistics of non-bankruptcy

| | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| ROE | 172 | -16.55 | 24.62 | .5783 | 3.13862 |
| ROA (EBITDA/A) | 171 | -1067.00 | 10.95 | -8.0035 | 82.34853 |
| ROA= EBIT/A | 171 | -1195.00 | 10.95 | -8.8219 | 92.08171 |
| ROS = EBIT/S | 170 | -120.25 | 8.89 | -1.6860 | 10.61126 |
| D/A | 171 | -19.72 | 1200.00 | 13.4850 | 98.94354 |
| Bank loans/A | 171 | .00 | 144.20 | 1.1495 | 11.02867 |
| Inventories*360/Sales | 170 | .00 | 1992.59 | 74.6836 | 241.72768 |
| receivables*30/Sales | 170 | -367.20 | 1239750.00 | 7801.3239 | 95103.71301 |
| A*360/Sales | 170 | -56.90 | 8404200.00 | 54086.8501 | 645363.60930 |
| current ratio | 172 | -3.53 | 11.86 | 1.1377 | 1.50510 |
| Quick ratio | 172 | -3.53 | 11.86 | .9055 | 1.46842 |
| Cash ratio | 172 | -5.97 | 8.78 | .1592 | .87207 |
| capital employed/Fixed assets | 135 | -1413.14 | 313.45 | -19.2113 | 138.57038 |
| (current assets-current liabilities)/current assets | 171 | -1199.00 | 50.90 | -11.9332 | 100.73708 |
| gross profit/sales | 170 | -71.00 | 1.15 | -.7512 | 6.21778 |
| Financial leverage | 172 | -98.17 | 408.15 | 8.0806 | 52.35056 |
| interests/ EBIT | 172 | -17.13 | 5.79 | -.1391 | 1.61080 |
| capital employed/Fixed assets | 135 | -1413.14 | 313.45 | -19.2113 | 138.57038 |

| | | | | | |
|---|---|---|---|---|---|
| current liabilities/A | 171 | -.28 | 1200.00 | 11.8495 | 98.15396 |
| D/E | 172 | -82.99 | 407.15 | 7.0037 | 51.70983 |
| Bank loans/E | 172 | -80.30 | 303.09 | 1.3130 | 25.42452 |